

1636



RECEIVED
JUG 13 2002
ENTER 1600/2900
600

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/424,458A

DATE: 08/08/2002

TIME: 11:56:57

Input Set : A:\13198.asc.txt

Output Set: N:\CRF4\08082002\I424458A.raw

#15

4 <110> APPLICANT: Hayward, Nicholas
5 Silins, Ginters
6 Grimmond, Sean
7 Gartside, Michael
8 Hancock, John
10 <120> TITLE OF INVENTION: THREE NOVEL GENES ENCODING A ZINC FINGER PROTEIN, A
11 GUANINE, NUCLEOTIDE EXCHANGE FACTOR AND A HEAT SHOCK
12 PROTEIN OR HEAT SHOCK BINDING PROTEIN
14 <130> FILE REFERENCE: 13198
16 <140> CURRENT APPLICATION NUMBER: 09/424,458A
17 <141> CURRENT FILING DATE: 2000-03-16
19 <150> PRIOR APPLICATION NUMBER: PCT/AU98/00380
20 <151> PRIOR FILING DATE: 1998-05-22
22 <160> NUMBER OF SEQ ID NOS: 125
24 <170> SOFTWARE: PatentIn Ver. 2.1
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 8
28 <212> TYPE: PRT
29 <213> ORGANISM: Artificial Sequence
31 <220> FEATURE:
32 <223> OTHER INFORMATION: Description of Artificial Sequence:Peptide repeat
33 motif in DnaJ homologues.
35 <220> FEATURE:
36 <221> NAME/KEY: UNSURE
37 <222> LOCATION: (2)..(3)
38 <223> OTHER INFORMATION: Xaa at position 2,3 can be any amino acid.
40 <220> FEATURE:
41 <221> NAME/KEY: UNSURE
42 <222> LOCATION: (5)
43 <223> OTHER INFORMATION: Xaa at position 5 can be any amino acid.
45 <220> FEATURE:
46 <221> NAME/KEY: UNSURE
47 <222> LOCATION: (7)
48 <223> OTHER INFORMATION: Xaa at position 7 can be any amio acid.
50 <400> SEQUENCE: 1
W--> 51 Cys Xaa Xaa Cys Xaa Gly Xaa Gly
52 1 5
54 <210> SEQ ID NO: 2
55 <211> LENGTH: 1242
56 <212> TYPE: DNA
57 <213> ORGANISM: Homo sapiens
59 <220> FEATURE:
60 <221> NAME/KEY: CDS

RAW SEQUENCE LISTING

DATE: 08/08/2002

PATENT APPLICATION: US/09/424,458A

TIME: 11:56:57

Input Set : A:\13198.asc.txt

Output Set: N:\CRF4\08082002\I424458A.raw

TECH CENTER 1600/2900

AUG 13 2002

RECEIVED

```

61 <222> LOCATION: (30)..(959)
63 <400> SEQUENCE: 2
64 tcagtaaaca cagagactgg ggatcgatc atg ggg ctt tgt aag tgc ccc aag 53
66 Met Gly Leu Cys Lys Cys Pro Lys
67 1 5
69 aga aag gtg acc aac ctg ttc tgc ttc gaa cat cgg gtc aac gtc tgc 101
71 Arg Lys Val Thr Asn Leu Phe Cys Phe Glu His Arg Val Asn Val Cys
72 10 15 20
74 gag cac tgc ctg gta gcc aat cac gcc aag tgc atc gtc cag tcc tac 149
76 Glu His Cys Leu Val Ala Asn His Ala Lys Cys Ile Val Gln Ser Tyr
77 25 30 35 40
79 ctg caa tgg ctc caa gat agc gac tac aac ccc aat tgc cgc ctg tgc 197
81 Leu Gln Trp Leu Gln Asp Ser Asp Tyr Asn Pro Asn Cys Arg Leu Cys
82 45 50 55
84 aac ata ccc ctg gcc agc cga gag acg acc cgc ctt gtc tgc tat gat 245
86 Asn Ile Pro Leu Ala Ser Arg Glu Thr Thr Arg Leu Val Cys Tyr Asp
87 60 65 70
89 ctc ttt cac tgg gcc tgc ctc aat gaa cgt gct gcc cag cta ccc cga 293
91 Leu Phe His Trp Ala Cys Leu Asn Glu Arg Ala Ala Gln Leu Pro Arg
92 75 80 85
94 aac acg gca cct gcc ggc tat cag tgc ccc agc tgc aat ggc ccc atc 341
96 Asn Thr Ala Pro Ala Gly Tyr Gln Cys Pro Ser Cys Asn Gly Pro Ile
97 90 95 100
99 ttc ccc cca acc aac ctg gct ggc ccc gtg gcc tcc gca ctg aga gag 389
101 Phe Pro Pro Thr Asn Leu Ala Gly Pro Val Ala Ser Ala Leu Arg Glu
102 105 110 115 120
104 aag ctg gcc aca gtc aac tgg gcc cgg gca gga ctg ggc ctc cct ctg 437
106 Lys Leu Ala Thr Val Asn Trp Ala Arg Ala Gly Leu Gly Leu Pro Leu
107 125 130 135
109 atc gat gag gtg gtg agc cca gag ccc gag ccc ctc aac acg tct gac 485
111 Ile Asp Glu Val Val Ser Pro Glu Pro Glu Pro Leu Asn Thr Ser Asp
112 140 145 150
114 ttc tct gac tgg tct agt ttt aat gcc agc agt acc cct gga cca gag 533
117 Phe Ser Asp Trp Ser Ser Phe Asn Ala Ser Ser Thr Pro Gly Pro Glu
118 155 160 165
120 gag gta gac agc gcc tct gct gcc cca gcc ttc tac agc cga gcc ccc 581
122 Glu Val Asp Ser Ala Ser Ala Ala Pro Ala Phe Tyr Ser Arg Ala Pro
123 170 175 180
125 cgg ccc cca gct tcc cca ggc cgg ccc gag cag cac aca gtg atc cac 629
127 Arg Pro Pro Ala Ser Pro Gly Arg Pro Glu Gln His Thr Val Ile His
128 185 190 195 200
130 atg ggc aat cct gag ccc ttg act cac gcc cct agg aag gtg tat gat 677
132 Met Gly Asn Pro Glu Pro Leu Thr His Ala Pro Arg Lys Val Tyr Asp
133 205 210 215
135 acg cgg gat gat gac cgg aca cca ggc ctc cat gga gac tgt gac gat 725
137 Thr Arg Asp Asp Asp Arg Thr Pro Gly Leu His Gly Asp Cys Asp Asp
138 220 225 230
140 gac aag tac cga cgt cgg ccg gcc ttg ggt tgg ctg gcc cgg ctg cta 773
142 Asp Lys Tyr Arg Arg Arg Pro Ala Leu Gly Trp Leu Ala Arg Leu Leu

```

RAW SEQUENCE LISTING

DATE: 08/08/2002

PATENT APPLICATION: US/09/424,458A

TIME: 11:56:57

Input Set : A:\13198.asc.txt

Output Set: N:\CRF4\08082002\I424458A.raw

```

143          235          240          245
145 agg agc cgg gct ggg tct cgg aag cgg ccg ctg acc ctg ctc cag cgg      821
147 Arg Ser Arg Ala Gly Ser Arg Lys Arg Pro Leu Thr Leu Leu Gln Arg
148          250          255          260
150 gcg ggg ctg ctg cta ctc ttg gga ctg ctg ggc ttc ctg gcc ctc ctt      869
152 Ala Gly Leu Leu Leu Leu Leu Gly Leu Leu Gly Phe Leu Ala Leu Leu
153 265          270          275          280
155 gcc ctc atg tct cgc cta ggc cgg gcc gca gct gac agc gat ccc aac      917
158 Ala Leu Met Ser Arg Leu Gly Arg Ala Ala Ala Asp Ser Asp Pro Asn
159          285          290          295
161 ctg gac cca ctc atg aac cct cac atc cgc gtg ggc ccc tcc      959
163 Leu Asp Pro Leu Met Asn Pro His Ile Arg Val Gly Pro Ser
164          300          305          310
166 tgagccccct tgcttgtggc taggccagcc taggatgtgg gttctgtgga ggagaggcgg 1019
169 ggtaatgggg aggctgaggg cacctcttca ctgccctct ccctcaagcc taagacacta 1079
172 agaccccaaga cccaaagcca agtccaccag agtggctcgc aggccaggcc tggagtcccc 1139
175 gtgggtcaag catttgtctt gacttgtctt ctcccgggtc tccagcctcc gacccctcgc 1199
178 cccatgaagg agctggcagg tggaaataaa caacaacttt att      1242
182 <210> SEQ ID NO: 3
183 <211> LENGTH: 310
184 <212> TYPE: PRT
185 <213> ORGANISM: Homo sapiens
187 <400> SEQUENCE: 3
188 Met Gly Leu Cys Lys Cys Pro Lys Arg Lys Val Thr Asn Leu Phe Cys
189 1          5          10          15
191 Phe Glu His Arg Val Asn Val Cys Glu His Cys Leu Val Ala Asn His
192          20          25          30
194 Ala Lys Cys Ile Val Gln Ser Tyr Leu Gln Trp Leu Gln Asp Ser Asp
195          35          40          45
197 Tyr Asn Pro Asn Cys Arg Leu Cys Asn Ile Pro Leu Ala Ser Arg Glu
198          50          55          60
200 Thr Thr Arg Leu Val Cys Tyr Asp Leu Phe His Trp Ala Cys Leu Asn
201 65          70          75          80
203 Glu Arg Ala Ala Gln Leu Pro Arg Asn Thr Ala Pro Ala Gly Tyr Gln
204          85          90          95
206 Cys Pro Ser Cys Asn Gly Pro Ile Phe Pro Pro Thr Asn Leu Ala Gly
207          100          105          110
210 Pro Val Ala Ser Ala Leu Arg Glu Lys Leu Ala Thr Val Asn Trp Ala
211          115          120          125
213 Arg Ala Gly Leu Gly Leu Pro Leu Ile Asp Glu Val Val Ser Pro Glu
214          130          135          140
216 Pro Glu Pro Leu Asn Thr Ser Asp Phe Ser Asp Trp Ser Ser Phe Asn
217 145          150          155          160
219 Ala Ser Ser Thr Pro Gly Pro Glu Glu Val Asp Ser Ala Ser Ala Ala
220          165          170          175
222 Pro Ala Phe Tyr Ser Arg Ala Pro Arg Pro Pro Ala Ser Pro Gly Arg
223          180          185          190
225 Pro Glu Gln His Thr Val Ile His Met Gly Asn Pro Glu Pro Leu Thr
226          195          200          205

```

RAW SEQUENCE LISTING

DATE: 08/08/2002

PATENT APPLICATION: US/09/424,458A

TIME: 11:56:57

Input Set : A:\13198.asc.txt

Output Set: N:\CRF4\08082002\I424458A.raw

```

229 His Ala Pro Arg Lys Val Tyr Asp Thr Arg Asp Asp Asp Arg Thr Pro
230      210      215      220
233 Gly Leu His Gly Asp Cys Asp Asp Asp Lys Tyr Arg Arg Arg Pro Ala
234 225      230      235      240
236 Leu Gly Trp Leu Ala Arg Leu Leu Arg Ser Arg Ala Gly Ser Arg Lys
237      245      250      255
239 Arg Pro Leu Thr Leu Leu Gln Arg Ala Gly Leu Leu Leu Leu Gly
240      260      265      270
242 Leu Leu Gly Phe Leu Ala Leu Leu Ala Leu Met Ser Arg Leu Gly Arg
243      275      280      285
245 Ala Ala Ala Asp Ser Asp Pro Asn Leu Asp Pro Leu Met Asn Pro His
246      290      295      300
248 Ile Arg Val Gly Pro Ser
249 305      310
254 <210> SEQ ID NO: 4
255 <211> LENGTH: 2415
256 <212> TYPE: DNA
257 <213> ORGANISM: Homo sapiens
259 <220> FEATURE:
260 <221> NAME/KEY: CDS
261 <222> LOCATION: (3)..(2186)
263 <400> SEQUENCE: 4
264 cg att tca ttc ctc gct ccc cac agg tcc ctc tcc cca aaa tat tcc      47
266   Ile Ser Phe Leu Ala Pro His Arg Ser Leu Ser Pro Lys Tyr Ser
267     1      5      10      15
269 cat ctt gtc cta gcc cat ccc cca gac tat ctc aag gac cag ctg tcc      95
271 His Leu Val Leu Ala His Pro Pro Asp Tyr Leu Lys Asp Gln Leu Ser
272      20      25      30
274 cca cgc ccc cga cct cca cta ggc ctg tgc cac ccg ctg cct gca gga      143
276 Pro Arg Pro Arg Pro Pro Leu Gly Leu Cys His Pro Leu Pro Ala Gly
277      35      40      45
279 aga cgc ccg gtc ccg ggc cgg gtt agc ccc atg gga acg cag cgc ctg      191
281 Arg Arg Pro Val Pro Gly Arg Val Ser Pro Met Gly Thr Gln Arg Leu
282      50      55      60
284 tgt ggc cgc ggg act caa ggc tgg cct ggc tca agt gaa cag cac gtc      239
286 Cys Gly Arg Gly Thr Gln Gly Trp Pro Gly Ser Ser Glu Gln His Val
287      65      70      75
289 cag gag gcg acc tcg tcc gcg ggt ttg cat tct ggg gtg gac gag ctg      287
291 Gln Glu Ala Thr Ser Ser Ala Gly Leu His Ser Gly Val Asp Glu Leu
292 80      85      90      95
294 ggg gtt cgg tcc gag ccc ggt ggg agg ctc ccg gag cgc agc ctg ggc      335
296 Gly Val Arg Ser Glu Pro Gly Gly Arg Leu Pro Glu Arg Ser Leu Gly
297      100      105      110
299 cca gcc cac ccc gcg ccg gcg gcc atg gca ggc acc ctg gac ctg gac      383
301 Pro Ala His Pro Ala Pro Ala Ala Met Ala Gly Thr Leu Asp Leu Asp
302      115      120      125
304 aag ggc tgc acg gtg gag gag ctg ctc cgc ggg tgc atc gaa gcc ttc      431
306 Lys Gly Cys Thr Val Glu Glu Leu Leu Arg Gly Cys Ile Glu Ala Phe
307      130      135      140

```

RAW SEQUENCE LISTING

DATE: 08/08/2002

PATENT APPLICATION: US/09/424,458A

TIME: 11:56:57

Input Set : A:\13198.asc.txt

Output Set: N:\CRF4\08082002\I424458A.raw

```

309 gat gac tcc ggg aag gtg cgg gac ccg cag ctg gtg cgc atg ttc ctc 479
311 Asp Asp Ser Gly Lys Val Arg Asp Pro Gln Leu Val Arg Met Phe Leu
312 145 150 155
314 atg atg cac ccc tgg tac atc ccc tcc tct cag ctg gcg gcc aag ctg 527
316 Met Met His Pro Trp Tyr Ile Pro Ser Ser Gln Leu Ala Ala Lys Leu
317 160 165 170 175
319 ctc cac atc tac caa caa tcc cgg aag gac aac tcc aat tcc ctg cag 575
321 Leu His Ile Tyr Gln Gln Ser Arg Lys Asp Asn Ser Asn Ser Leu Gln
322 180 185 190
324 gtg aaa acg tgc cac ctg gtc agg tac tgg atc tcc gcc ttc cca gcg 623
326 Val Lys Thr Cys His Leu Val Arg Tyr Trp Ile Ser Ala Phe Pro Ala
327 195 200 205
329 gag ttt gac ttg aac ccg gag ttg gct gag cag atc aag gag ctg aag 671
331 Glu Phe Asp Leu Asn Pro Glu Leu Ala Glu Gln Ile Lys Glu Leu Lys
332 210 215 220
334 gct ctg cta gac caa gaa ggg aac cga cgg cac agc agc cta atc gac 719
336 Ala Leu Leu Asp Gln Glu Gly Asn Arg Arg His Ser Ser Leu Ile Asp
337 225 230 235
339 ata gac agc gtc cct acc tac aag tgg aag cgg cag gtg act cag cgg 767
341 Ile Asp Ser Val Pro Thr Tyr Lys Trp Lys Arg Gln Val Thr Gln Arg
342 240 245 250 255
344 aac cct gtg gga cag aaa aag cgc aag atg tcc ctg ttg ttt gac cac 815
346 Asn Pro Val Gly Gln Lys Lys Arg Lys Met Ser Leu Leu Phe Asp His
347 260 265 270
349 ctg gag ccc atg gag ctg gcg gag cat ctc acc tac ttg gag tat cgc 863
351 Leu Glu Pro Met Glu Leu Ala Glu His Leu Thr Tyr Leu Glu Tyr Arg
352 275 280 285
354 tcc ttc tgc aag atc ctg ttt cag gac tat cac agt ttc gtg act cat 911
356 Ser Phe Cys Lys Ile Leu Phe Gln Asp Tyr His Ser Phe Val Thr His
357 290 295 300
359 ggc tgc act gtg gac aac ccc gtc ctg gag cgg ttc atc tcc ctc ttc 959
361 Gly Cys Thr Val Asp Asn Pro Val Leu Glu Arg Phe Ile Ser Leu Phe
362 305 310 315
364 aac agc gtc tca cag tgg gtg cag ctc atg atc ctc agc aaa ccc aca 1007
366 Asn Ser Val Ser Gln Trp Val Gln Leu Met Ile Leu Ser Lys Pro Thr
367 320 325 330 335
369 gcc ccg cag cgg gcc ctg gtc atc aca cac ttt gtc cac gtg gcg gag 1055
371 Ala Pro Gln Arg Ala Leu Val Ile Thr His Phe Val His Val Ala Glu
372 340 345 350
374 aag ctg cta cag ctg cag aac ttc aac acg ctg atg gca gtg gtc ggg 1103
376 Lys Leu Leu Gln Leu Gln Asn Phe Asn Thr Leu Met Ala Val Val Gly
377 355 360 365
379 ggc ctg agc cac agc tcc atc tcc cgc ctc aag gag acc cac agc cac 1151
381 Gly Leu Ser His Ser Ser Ile Ser Arg Leu Lys Glu Thr His Ser His
382 370 375 380
384 gtt agc cct gag acc atc aag ctc tgg gag ggt ctc acg gaa cta gtg 1199
386 Val Ser Pro Glu Thr Ile Lys Leu Trp Glu Gly Leu Thr Glu Leu Val
387 385 390 395
389 acg gcg aca ggc aac tat ggc aac tac cgg cgt cgg ctg gca gcc tgt 1247

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/424,458A

DATE: 08/08/2002
TIME: 11:56:58

Input Set : A:\13198.asc.txt
Output Set: N:\CRF4\08082002\I424458A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 2,3,5,7
Seq#:49; Xaa Pos. 10,21,55,77,87
Seq#:74; Xaa Pos. 2,3,5,6,7,8,10,11,13,14,15,16,18,19,21,22,23,24,25,26,27
Seq#:74; Xaa Pos. 28,29,30,31,32,33,34,35,36,37,39,40,42,43,44,45,46,47,48
Seq#:74; Xaa Pos. 49,50,51,52,53,54,55,56,57,58,59,61,62,64,65,66,67,68,69
Seq#:74; Xaa Pos. 70,71,72,73,74,75,76,77,78,79,80,81,83,84
Seq#:75; Xaa Pos. 2,3,4,5,6,7,9,10,11,12,13,14,16,17,18,19,20,21,23,24,25
Seq#:75; Xaa Pos. 26,27,28
Seq#:76; Xaa Pos. 2,3,4,5,6,7,9,11,12,13,14,15,16,18,20,21,22,23,24,25
Seq#:113; Xaa Pos. 4,35
Seq#:121; Xaa Pos. 11
Seq#:122; Xaa Pos. 35

VERIFICATION SUMMARY

DATE: 08/08/2002

PATENT APPLICATION: US/09/424,458A

TIME: 11:56:58

Input Set : A:\13198.asc.txt

Output Set: N:\CRF4\08082002\I424458A.raw

L:51 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:0
L:1768 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49 after pos.:0
L:1771 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49 after pos.:16
L:1777 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49 after pos.:48
L:1780 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49 after pos.:64
L:1783 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49 after pos.:80
L:2173 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74 after pos.:0
L:2176 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74 after pos.:16
L:2179 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74 after pos.:32
L:2182 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74 after pos.:48
L:2185 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74 after pos.:64
L:2188 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74 after pos.:80
L:2222 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:75 after pos.:0
L:2225 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:75 after pos.:16
L:2264 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:76 after pos.:0
L:2267 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:76 after pos.:16
L:3396 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:113 after pos.:0
L:3402 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:113 after pos.:32
L:3533 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:121 after pos.:0
L:3557 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:122 after pos.:32